

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-305571

(43)Date of publication of application : 28.11.1997

(51)Int.Cl.

G06F 17/00
G06F 17/60

(21)Application number : 08-116004

(71)Applicant : HITACHI LTD

(22)Date of filing : 10.05.1996

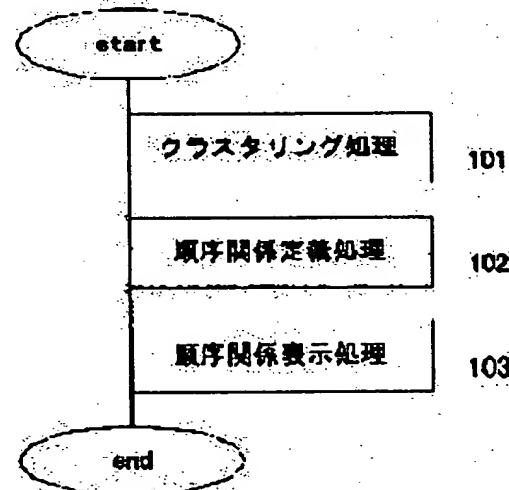
(72)Inventor : ASHIDA HITOSHI
MAEDA AKIRA
ITO YUKIYASU

(54) TIME SERIES DATA PROCESSING METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To accelerate processing by reducing the number of combination of merchandise to define the relation of order by automatically determining the merchandise (service) vending order of a customer based on applied time series data.

SOLUTION: In the case of order relation definition processing 101, clustering processing is performed for classifying the merchandise into the plural groups of similar customers and concerning respective generated classes, the relation of order to purchase the respective articles is defined. In order relation definition processing 102, when the order relation of purchase of two articles is recognized from the counted number of events, the average value of time difference and the standard deviation are found and the order relation of merchandise in the respective classes is displayed. In order relation display processing 103, the drawing of order relation between the merchandise designated by a user and the other merchandise in the same class is prepared and graphically displayed. In this case, when any one merchandise is designated, the order relation with the other merchandise in the same class is uniquely determined by referring to an order relation definition table generated by the processing 102.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平9-305571

(43) 公開日 平成9年(1997)11月28日

(51) Int.Cl.[®]
G 0 6 F 17/00
17/60

識別記号 庁内整理番号
17/00

F I
G 0 6 F 15/20
15/21

技術表示箇所
N
3 1 0 Z

審査請求 未請求 請求項の数5 OL (全7頁)

(21) 出願番号 特願平8-116004

(22) 出願日 平成8年(1996)5月10日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 芦田 仁史

神奈川県川崎市麻生区玉禅寺1099番地株式

会社日立製作所システム開発研究所内

(72) 発明者 前田 章

神奈川県川崎市麻生区玉禅寺1099番地株式

会社日立製作所システム開発研究所内

(72) 発明者 伊藤 幸康

神奈川県横浜市戸塚区戸塚町5030番地株式

会社日立製作所ソフトウェア開発本部内

(74) 代理人 弁理士 小川 勝男

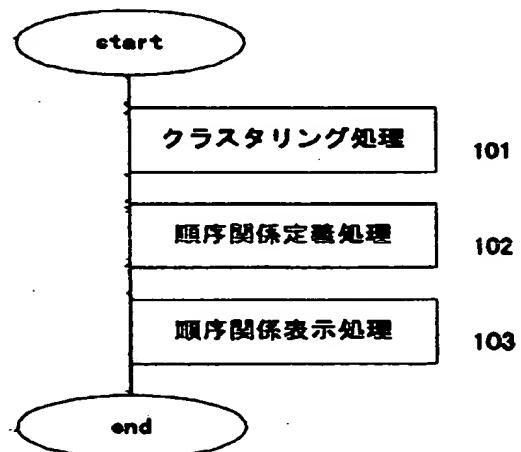
(54) 【発明の名称】 時系列データ処理方法

(57) 【要約】

【課題】 处理日時などの時間情報、顧客No.などの顧客を特定する情報、商品名などの商品またはサービスを特定する情報を最低限含むレコードからなる時系列データから、顧客が商品(サービス)を購買する順序を自動的に決定する。

【解決手段】 处理日時などの時間情報、顧客No.などの顧客を特定する情報、商品名などの商品またはサービスを特定する情報を最低限含むレコードからなる時系列データを対象とし、商品をその購買顧客が類似した複数のグループに分類するクラスタリング処理101と、該クラスタリング処理により生成された各クラス内の商品間の購買される順序関係を定義する順序関係定義処理102と、ユーザが指定した商品と、同一クラス内の他の商品との順序関係を表示する順序関係表示処理103から構成される。

図1



1

【特許請求の範囲】

【請求項1】処理日時などの時間情報、顧客No.などの顧客を特定する情報、商品名などの商品またはサービスを特定する情報を最低限含むレコードから成る時系列データを処理する方法において、該時系列データに基づき、顧客の商品（サービス）購買順序を自動的に決定する処理を備えたことを特徴とする時系列データ処理方法。

【請求項2】請求項1に記載の時系列データ処理方法において、商品をその購買顧客が類似した複数のグループに分類するクラスタリング処理と、該クラスタリング処理により生成された各クラス内の商品間の購買される順序関係を定義する順序関係定義処理と、ユーザが指定した商品と、同一クラス内の他の商品との順序関係を表示する順序関係表示処理を備えたことを特徴とする時系列データ処理方法。

【請求項3】請求項2に記載の時系列データ処理方法において、上記クラスタリング処理は、任意の2つの商品の距離定義として、2つの商品を購入した顧客層の違いを利用することを特徴とする時系列データ処理方法。

【請求項4】請求項2に記載の時系列データ処理方法において、上記グループ内順序関係定義処理は、クラスを選択する処理と、各クラス内で2つの商品の組合せを全て生成する処理と、2つの商品の購買が同時におこなわれたレコード数および順次購買されたレコード数をカウントする処理と、該レコード数に基づき、商品間の購買の順序関係を決定する処理と、順序関係が成立した商品間の時間差の平均値と標準偏差を求める処理と、クラス毎に商品間の順序関係まとめた順序関係定義表を作成する処理を備えたことを特徴とする時系列データ処理方法。

【請求項5】請求項2に記載の時系列データ処理方法において、上記順序関係表示処理は、上記順序関係定義表を参照し、ユーザが指定した商品と、同一グループ内の他の商品との順序関係図を作成し、グラフィック表示する処理を備えたことを特徴とする時系列データ処理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、販売情報などの時系列データから、顧客の商品購買の順序関係を自動的に求める方法に関する。

【0002】

【従来の技術】近年、ディスクの大容量化および低価格化、CPUの高速化および並列分散処理の実用化、またはデータベースシステムのダウンサイ징への適用が進み、大量データの蓄積が容易かつ安価になった。プラントにおける、品質管理データ、操業データ、ビジネス分野における顧客データ、POSデータ、トランザクションデータなどである。また、これらのデータの多く

2

は、常時追加、更新が行われる時系列データである。

【0003】これらの時系列データを解析する手法としては、重回帰分析、相関分析などの多変量解析の手法が広く知られている。（従来技術1）

また、時系列データをフィルタリングによりスペクトルに変換し、データの特徴を抽出し、線形予測や、パターンマッチングをおこなう方法が、「パターン認識と学習のアルゴリズム」（文一総合出版、1990）なる文献に解説されている。（従来技術2）

10 また、POSデータに含まれる特徴を抽出する技術としては、同時に購買されている商品の組を抽出し、さらに、それらの商品の因果関係を求めるAssociation Rulesと呼ばれる技術がある。AAAI Workshop on Knowledge Discovery in DatabasesのHeikki MannilaによるEfficient Algorithms for Discovering Association Rules (p 181-192) になる文献に解説がある。（従来技術3）

【0004】

【発明が解決しようとする課題】上記従来技術1は、複数の系列の類似性を比較するものであり、1つの系列の中に現れる特徴を抽出するものではない。

【0005】上記従来技術2は、一つの系列の中の特徴を抽出する技術ではあるが、音声認識などの、連続的に現れるデータに適しており、トランザクションデータ中の特徴を抽出するのは困難である。

【0006】上記従来技術3で述べられているAssociation Rulesは、マーケットバスケットと呼ばれる、顧客が同時に購入する商品の組合せとそれらの因果関係を求めるものである。しかし、異なる時間や場所で購入された商品の関連を求めるものではない。

【0007】上記従来技術1、2、3の問題点に鑑み、本発明は、時系列データ中に含まれる特徴を事例発生の順序関係の形で抽出することを第1の目的とする。

【0008】さらに、抽出した特徴を直観的に理解できる形式で表示することを第2の目的とする。

【0009】

【課題を解決するための手段】上記の目的を達成するため、本発明の時系列データ処理方法では、時系列データに基づき、顧客の商品購買順序を自動的に決定する処理を備えている。

【0010】特に本発明の望ましい態様では、商品をその購買顧客が類似した複数のグループに分類するクラスタリング処理と、該クラスタリング処理により生成された各クラス内の商品間の購買される順序関係を定義する順序関係定義処理と、ユーザが指定した商品と、同一クラス内の他の商品との順序関係を表示する順序関係表示処理を備えている。

【0011】具体的には、上記クラスタリング処理は、任意の2つの商品の距離定義として、2つの商品を購入した顧客層の違いを利用することを特徴とする。

【0012】上記グループ内順序関係定義処理は、クラスを選択する処理と、各クラス内で2つの商品の組合せを全て生成する処理と、2つの商品の購買が同時に起こなわれたレコード数および順次購買されたレコード数をカウントする処理と、該レコード数に基づき、商品間の購買の順序関係を決定する処理と、順序関係が成立した商品間の時間差の平均値と標準偏差を求める処理と、クラス毎に商品間の順序関係まとめた順序関係定義表を作成する処理を備えている。

【0013】上記順序関係表示処理は、上記順序関係定義表を参照し、ユーザが指定した商品と、同一グループ内の他の商品との順序関係図を作成し、グラフィック表示する処理を備えている。

【0014】

【発明の実施の形態】以下、図面を用いて本発明の実施の形態について説明する。

【0015】図1は、本発明の対象である時系列データ処理方法のフローチャートである。この方法を用いて対象とする時系列データから顧客の商品購買順序を求める。

任意の2つの商品A、B間の距離(類似度)は、

$$\text{距離}(A, B) = \frac{(A \text{ を購入した顧客数}) + (B \text{ を購入した顧客数}) - (A, B \text{ を購入した顧客数})}{\dots \dots \dots \text{ (数1)}}$$

【0020】数1は任意の2商品A、Bの距離を示す。A、Bをいずれも購入した顧客数の、どちらか一方でも購入した顧客数に対する割合を1から引いたものである。商品A、Bを購入する顧客が全て同じなら最小値0となり、全て異なれば最大値1となる。2つの商品いずれも購入している顧客数の、いずれか一方でも購入している顧客数に対する割合が大きくなるほど、距離は0に近くなる。例えば、クラスタリングの結果は、図3に示すような形式にまとめられる。

【0021】なお、このクラスタリング処理は必ずしも実行する必要はない。この処理をスキップした場合には、クラス数が一つとして後の処理を実行する。

【0022】次に順序関係定義処理102の詳細を図4に示す。図4に示す通り、順序関係定義処理は、ユーザがクラスを選択するクラス選択処理401、任意の2商品(例えば、A、Bとする)を順次選択し、処理404以降を繰り返す商品選択処理402、2つの商品が同時に購入されている事例数、Aが購入された後にBが購入されている事例数、Bが購入された後にAが購入されている事例数をカウントする事例カウント処理404、処理404でカウントした事例数から、2つの商品の順序関係を決定する商品関係決定処理405、2つの商品の購買に順序関係が認められた場合に、時間差の平均値と標準偏差を求める時間差算出処理406、各クラス内の商品間の順序関係を表示する順序関係定義表生成処理4

【0016】図1において101は、商品をその購買顧客が類似した複数のグループに分類するクラスタリング処理、生成された各クラスにおいて、各商品間の購買される順序関係を定義する順序関係定義処理、103は、同一クラス内の他の商品との順序関係を表示する順序関係表示処理である。対象とする時系列データは例えば、図2に示す形式で与えられる。

【0017】次に図1の各処理について詳しく説明する。まず、クラスタリング処理101について説明する。クラスタリング処理101では、ユーザが指定した商品を、ユーザが指定したクラス数に分類する。クラスタリング手法としては、一般的に知られている最長距離法などを用いる。

【0018】ここで、クラスタリングに用いる距離の定義式を以下の(数1)に示す。

【0019】

【数1】

03から成る。

【0023】商品関係決定処理405では、Aを購入した後にBを購入した顧客数が、A、B同時に購入した顧客数以上であり、かつAを購入した後にBを購入した顧客数が、Bを購入した後にAを購入した顧客数よりも多い場合は、AがBに先行することを示す(A→B)型と判断し、Bを購入した後にAを購入した顧客数が、A、B同時に購入した顧客数以上であり、かつBを購入した後にAを購入した顧客数が、Aを購入した後にBを購入した顧客数よりも多い場合は、BがAに先行することを示す(B→A)型と判断し、A、B同時に購入した顧客数が、Aを購入した後にBを購入した顧客数、Bを購入した後にAを購入した顧客数のいずれよりも多い場合は、AとBは同時に購入されたことを示す同時購入型と判断し、Aを購入した後にBを購入した顧客数が、Bを購入した後にAを購入した顧客数と等しく、かつ、Aを購入した後にBを購入した顧客数が、A、B同時に購入した顧客数よりも多い場合は、判別が不可能であることを示す判別不能型と判断し、判断した結果を出力する。

【0024】処理405において、同時とは、厳密に同時刻である必要はなく、ユーザが定義可能とする。例えば、同日のレコードを全て同時に定義することもできる。

【0025】時間差算出処理406では、商品関係決定処理405において、(A→B)型、あるいは(B→

A) 型と判断された商品の組合せにおいて、購買された時間差の平均値と標準偏差を求める。

【0026】対象とするデータは、

- a) 商品A, Bを購入した全顧客
- b) 商品A, Bを購入した顧客の中で、処理405において判断された順序関係を満たす全顧客

のいずれかをユーザが選択できる。

【0027】順序関係定義表生成処理403では、処理402, 404, 405, 406において決定した順序関係の結果をまとめた順序関係定義表を作成する。商品A, B, C, Dが同一クラスにカテゴリ化された場合に作成される表の例を図5にまとめる。図5に示す表の各セルには、商品購買の順序関係、順序関係が成立した場合には、その時間差の平均値と標準偏差を記入している。ここで、順序関係の右隣の括弧内に記述されている数値は、(定義した順序関係が成立している顧客数) / (その2つの商品を購買している顧客数) である。

【0028】次に順序関係表示処理103について説明する。順序関係表示処理103は、ユーザが指定した商品と、同一クラス内の他の商品との順序関係図を作成し、グラフィック表示する。商品を1つ指定すれば、同一クラス内の他の商品との順序関係は処理102において生成される順序関係定義表を参照することにより一意に決定する。例えば、図5に示す順序関係定義表が成立している場合に、商品Aを選択した場合の順序関係図を図6に示す。図6において、Aに先行して購入されたDはAの左側、Aの後に購入されたBはAの右側におかれて、矢印で結ばれている。矢印は時間の推移を表す。矢印の下には、順序関係定義表に記述されている値が記入されている。Aと同時に購入されたCはAの真下に記述され、二重線で結ばれている。時間の推移が存在しないことを表している。二重線の右側には順序関係定義表に記述されている値が記入されている。

【0029】以上の時系列データ処理方法の具体的利用例である。販売戦略立案支援システムについて説明する。ある会員制のスーパーマーケットでは、図2に示す形式の時系列データが蓄積されている。このデータを用いて、商品配列案、ダイレクトマーケティング戦略案、需要量予測案を自動的に作成する。図7に販売戦略立案支援システムの構成図を示す。販売戦略立案支援システムは、本発明の時系列データ処理702と、商品配列決定処理707、ダイレクトマーケティング支援処理708、需要量予測方式決定支援処理709により構成される。時系列データ処理702は、時系列データ701、ユーザ指定情報703を入力とし、クラスタリング結果表704、順序関係定義表705、順序関係図706を出力する。商品配列決定処理707は、クラスタリング結果704、順序関係定義表705およびユーザ指定情報703を入力とし、商品配列案710を出力する。ダイレクトマーケティング支援処理708は、順序関係定

義表705およびユーザ指定情報703を入力とし、ダイレクトマーケティング戦略案711を出力する。需要量予測方式決定支援処理709は、順序関係定義表709、ユーザ指定情報703を入力とし、需要量予測方式案712を出力する。また、ユーザは、順序関係図706を参照し、ユーザ指定情報703を決定する。

【0030】まず、商品配列案710を出力するプロセスを具体的に説明する。商品配列案710を出力するには、まず、時系列データ処理702により、クラスタリング結果704、順序関係定義表705を出力する。

【0031】ここで入力するユーザ指定情報703は、

- ・対象商品

- ・同時の定義

であり、ここでは、対象商品を全商品、同時の定義として同日を選択する。クラスタリング結果704は図3に示す形式、順序関係定義表705は図5に示す形式で表される。ただし、図5はクラス1に関する順序関係定義表であり、クラスの数だけ、同様な表が存在する。

【0032】商品配列決定処理では、ユーザ指定情報703により、以下に述べる3手法の中から1つを選択する。

【0033】(1) クラス参照

クラスタリング結果704のみ参照し、同一クラスに含まれる商品は近傍に配置する。この例だと、出力される商品配置案710は、

- i) 商品A, B, C, Dを近傍に配置
- ii) 商品E, G, . . . を近傍に配置

などである。

【0034】(2) 同時購買参照

30 各クラスの順序関係定義表を参照し、同時に購買される商品を近傍に配置することとする。この例だと、出力される商品配置案710は、i) 商品A, Cを近傍に配置などである。

【0035】(3) ユーザ指定

順序関係図を参照し、ユーザが近傍に配置する商品を決定する。例えば、図6に示す順序関係図を参照し、70%の顧客が商品D, Aの両方を購入しているし、時間差も15日しかないので、近傍に配置する。と決定する。

【0036】次にダイレクトマーケティング戦略71140を出力するプロセスを具体的に説明する。まず、時系列データ処理702により、クラスタリング結果704、順序関係定義表705を出力するところまでは、商品配列案710を出力する時と同様である。

【0037】ダイレクトマーケティング支援処理708では、ユーザ指定情報703を参照し、ダイレクトマーケティング戦略案を作成する。ここで指定されるユーザ指定情報703は、

- ・事例数の最小値

- ・事例の割合の最小値

- ・時間差の標準偏差の最大値

・ダイレクトメール発送のタイミング

等の中からユーザが指定可能である。例えば、事例数の最小値=30、事例の割合の最小値=65%、時間差の標準偏差の最大値=10日、ダイレクトメール発送のタイミング=時間差の平均値-2日と指定すれば、出力されるダイレクトマーケティング戦略案711は、「商品Dを購買した顧客に対し、購買日の13日後にダイレクトメールを送付」などである。

【0038】次に需要量予測方式案を出力するプロセスを具体的に説明する。まず、時系列データ処理702により、クラスタリング結果704、順序関係定義表705を出力するところまでは、商品配列案710を出力する時と同様である。

【0039】需要量予測方式決定支援処理709では、ユーザ指定情報703を参照し、需要量予測方式案を出力する。ここで指定されるユーザ指定情報703は、

- ・事例数の最小値
- ・事例の割合の最小値
- ・時間差の標準偏差の最大値

等の中からユーザが指定可能である。例えば、事例数の最小値=30、事例の割合の最小値=65%、時間差の標準偏差の最大値=10日と指定すれば、出力される需要量予測方式案は、

「商品Aの需要量予測値=15日前の商品Dの販売量×0.7」

となる。

【0040】以上、本発明の時系列データ処理方法を販売戦略立案システムに適用した例を説明した。この例は、時系列データとして販売実績データを利用した例である。この例では、各商品間の順序関係を求めたが、単体の商品である必要は無く、オーディオ製品、家電製品等、商品のグループも対象にできる。また、本発明が対象とするデータは、販売実績データに限定されない。本発明の時系列データ処理方法は、時間的順序関係が含まれるデータならば、例えば、プラントにおける商品管理データにも適用可能である。

*【0041】

【発明の効果】本発明の時系列データ処理方法により、与えられる時系列データに基づき、顧客の商品（サービス）販売順序を自動的に決定することが可能になる。

【0042】本発明によれば、あらかじめ商品をその購買顧客が類似した複数のグループに分類することにより、順序関係を定義する商品の組み合わせの数を削減し、処理を高速化できる。

【0043】本発明によれば、商品間の距離定義として、2つの商品を購入した顧客層の違いを利用することにより、顧客が類似した商品のグループを生成できる。

【0044】本発明によれば、商品間の購買の順序関係を定義した順序関係定義表を作成することにより、ユーザがある商品を指定すれば、その商品と他の商品の購買の順序関係を即座に表示できる。

【0045】本発明によれば、ユーザが指定した商品と他の商品の購買の順序関係をグラフィック表示することにより、ユーザは購買の順序関係を直感的に理解できる。

【0046】また、本発明の時系列データ処理方法を利用すれば、商品配置案、ダイレクトメール戦略案、需要量予測方法案を容易に生成でき、販売戦略立案システム等を構築できる。

【図面の簡単な説明】

【図1】本発明のフローチャートである。

【図2】本発明の処理の対象である時系列データの例である。

【図3】商品をクラスタリングした結果の例である。

【図4】商品の順序関係を定義する順序関係定義処理102のパド図である。

【図5】順序関係定義処理102の出力である順序関係定義表の例である。

【図6】順序関係表示処理103の出力である順序関係図の例である。

【図7】本発明の時系列データ処理方法を適用した販売戦略立案システムの構成図である。

【図2】

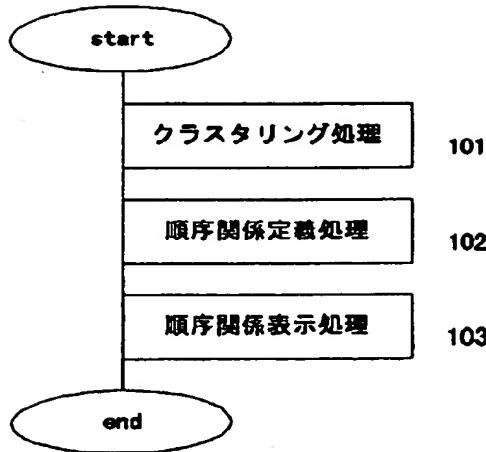
図2				
レコード番.	顧客番.	処理時間	商品名	...
0000001	01322300	1995/1/10 10:00	C	...
0000002	02537629	1995/1/10 10:25	F	...
...

【図3】

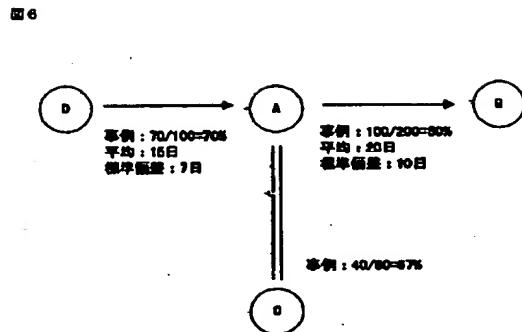
図3	
クラス名	商品名
クラス1	A, B, C, D
クラス2	E, G, ...
クラス3	F, ...
:	:
クラスn	H, ...

【図1】

図1

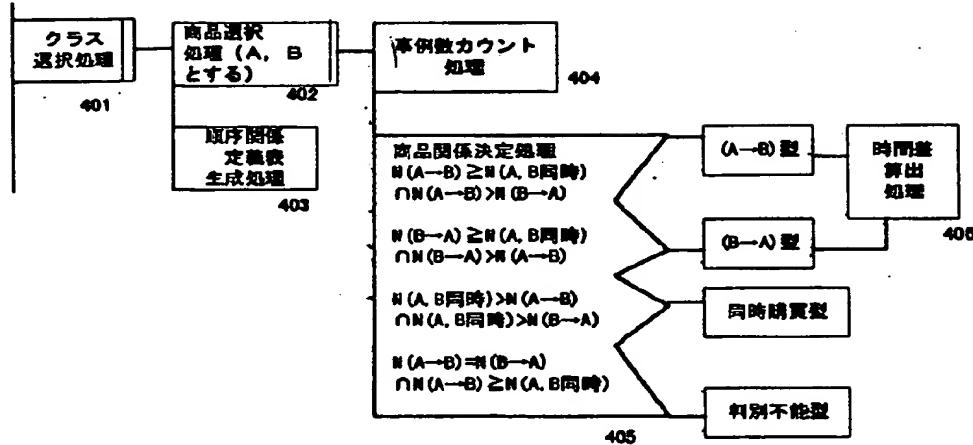


【図6】



【図4】

図4



* $N(A \rightarrow B)$: Aを購入した後にBを購入した顧客数
 $N(B \rightarrow A)$: Bを購入した後にAを購入した顧客数
 $N(A, B\text{同時})$: A, Bを同時に購入した顧客数

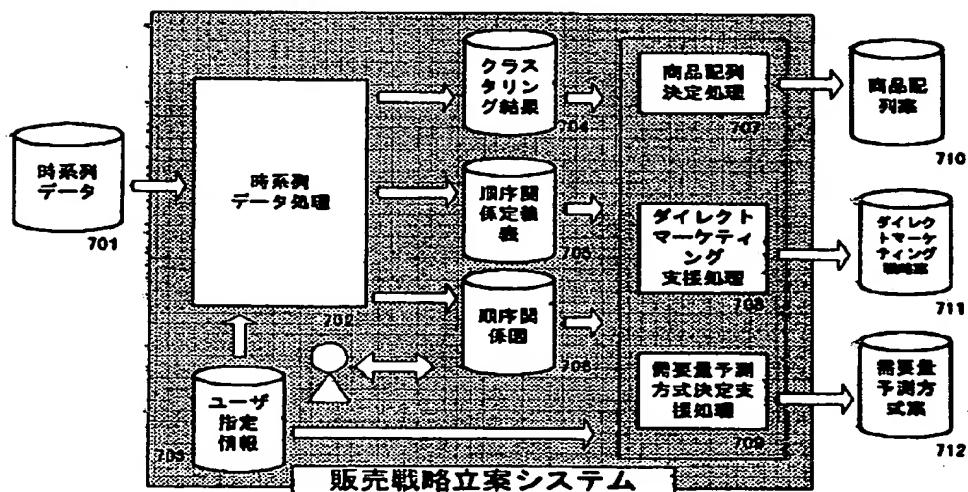
【図5】

図5

	商品A	商品B	商品C	商品D
商品A		A→B (100/200) 平均: 20日 標準偏差: 10日	同時 (40/60)	D→A (70/100) 平均: 15日 標準偏差: 7日
商品B	A→B (100/200) 平均: 20日 標準偏差: 10日		C→B (50/100) 平均: 18日 標準偏差: 10日	D→B (90/200) 平均: 60日 標準偏差: 20日
商品C	同時 (40/60)	C→B (50/100) 平均: 18日 標準偏差: 10日		D→C (80/110) 平均: 18日 標準偏差: 16日
商品D	D→A (70/100) 平均: 15日 標準偏差: 7日	D→B (90/200) 平均: 60日 標準偏差: 20日	D→C (80/110) 平均: 18日 標準偏差: 16日	

【図7】

図7



**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.